

Zentralitätsanomalien und Netzwerkstruktur: ein Plädoyer für einen "engeren" Netzwerkbegriff und ein community-orientiertes Zentralitätsmodell

Mutschke, Peter

Veröffentlichungsversion / Published Version

Sammelwerksbeitrag / collection article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Mutschke, P. (2008). Zentralitätsanomalien und Netzwerkstruktur: ein Plädoyer für einen "engeren" Netzwerkbegriff und ein community-orientiertes Zentralitätsmodell. In C. Stegbauer (Hrsg.), *Netzwerkanalyse und Netzwerktheorie: ein neues Paradigma in den Sozialwissenschaften* (S. 261-272). Wiesbaden: VS Verl. für Sozialwiss. https://doi.org/10.1007/978-3-531-91107-6_20

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

gesis
Leibniz-Institut
für Sozialwissenschaften

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Mitglied der
Leibniz
Leibniz-Gemeinschaft

Zentralitätsanomalien und Netzwerkstruktur. Ein Plädoyer für einen „engeren“ Netzwerkbegriff und ein community-orientiertes Zentralitätsmodell¹

Peter Mutschke

Zentralitätsanalysen führen in der Praxis gelegentlich zu erheblichen Plausibilitätsproblemen. Dies gilt insbesondere für das Betweenness-Konzept. Empirisch lässt sich zeigen, dass Plausibilitätsverluste oftmals mit Inkonsistenzen zwischen den Resultaten verschiedener Zentralitätsmaße wie Degree, Closeness und Betweenness einhergehen. Der Beitrag zeigt, dass Statusinkonsistenzen auf die mangelnde Berücksichtigung von Netzwerkstruktur in Standardzentralitätsmaßen der Netzwerkanalyse zurückgeführt werden können. Für solche strukturinduzierten Verzerrungen von Zentralitätswerten wird der Begriff der *Zentralitätsanomalie* eingeführt. Das Paper stellt ein Metamodell von Zentralität vor, das Zentralitätsanomalien durch Verwendung eines „engeren“, an der Community-Struktur sozialer Netzwerke orientierten Netzwerkbegriffs zu vermeiden sucht.

1 Einleitung: Zentralität und Ranking

In der Netzwerkanalyse spielt das Zentralitätskonzept eine überaus große Rolle. Es liegt daher nahe, Zentralität auch in Rankingkontexten wie Information Retrieval und Forschungsevaluation zu nutzen. Netzwerkanalytische Methoden wurden hier bisher jedoch nur in sehr eingeschränktem Maße eingesetzt. Eine berühmte Ausnahme ist die Internet-Suchmaschine Google, welche den Verlinkungsgrad von Web-Seiten für das Ranking nutzt (Brin/ Page 1998). Das Beispiel Google demonstriert bereits das Potential der Netzwerkanalyse für Suchmaschinen, schöpft es jedoch bei weitem nicht aus. Weitere bislang recht wenig genutzte Anwendungsfelder für Rankings bieten wissenschaftliche Kooperationsnetzwerke. Scientometrische Studien zeigen, dass diese Netzwerke einen enormen Einfluss auf das lokale, nationale und internationale Forschungsgeschehen haben (vgl. Beaver 2004). Wissenschaftliche Kooperationsnetzwerke bergen offensichtlich ein erhebliches Innovationspotential sowohl für die Forschungsevaluation als auch für die Suche in wissenschaftlichen Literaturdatenbanken.

Am Informationszentrum Sozialwissenschaften der GESIS² wurde ein Modell entwickelt, dass das netzwerkanalytische Konzept der Zentralität für die Suche nach Experten und das Ranking von Rechercheergebnissen auf der Basis von wissenschaftlichen Literatur- und Forschungsprojektdatenbanken verwendet (Mutschke 2004). Grundlage hierfür sind soziale Netzwerke, die sich durch Kooperation wissenschaftlicher Autoren konstituieren (im Folgenden kurz *Autorennetzwerke* genannt). Die Grundannahme für ein solches Ran-

¹ Danksagung: Ich danke Wolfgang Sodeur für einige sehr wertvolle Hinweise.

² <http://www.gesis.org/IZ/index.htm>.

king ist, dass strategisch herausragende Positionen in Autorennetzwerken auch qualitativ hochwertige Informationslieferanten für die Community repräsentieren. Das Modell identifiziert zentrale Autoren innerhalb einer Dokumentenkollektion bzw. Treffermenge, indem es auf der Basis der in den Dokumenten enthaltenen Informationen zu Kooperationsbeziehungen zwischen Autoren (z.B. Koautorenschaften) per Cword-Analyse ein Autorennetzwerk generiert und die Zentralität der Autoren in diesem Netzwerk ermittelt. Die Dokumente der Ergebnismenge werden dann mit dem Zentralitätswert ihres jeweils zentralsten Autors gewichtet und nach diesem Rankingwert absteigend sortiert ausgegeben. An der Spitze der Liste stehen somit Dokumente, deren Autoren eine hohe Zentralität im Autorennetzwerk aufweisen.

Dieses Verfahren wird bereits erfolgreich in Informationssystemen der GESIS wie infoconnex³, aber auch für Evaluationskontexte eingesetzt (Mutschke/ Stahl 2005). Die Standard-Rankingmethode dieses Verfahrens ist Betweenness-Zentralität. Erste heuristische Retrievaltests demonstrieren, dass die Anwendung von Zentralitätsanalysen für derartige Rankingkontexte ein erhebliches Precision-Potential hat (Mutschke 2004)⁴. Gelegentlich gibt es allerdings aus Nutzersicht auch unerwartete Plausibilitätsprobleme, die teilweise jedoch so erheblich sind, dass sie den Aussagewert der Zentralitätsanalyse insgesamt in Frage stellen. Das Paper diskutiert diese Problematik anhand eines exemplarischen Beispiels aus einer aktuellen Evaluationsstudie und versucht, vor dem Hintergrund einer analytischen Betrachtung des Problems und seiner Ursachen einerseits und aktuellen Erkenntnissen der Theoretischen Physik über die strukturelle Beschaffenheit sozialer Netzwerke andererseits, einen möglichen Ausweg aufzuzeigen. Doch zunächst noch einmal zurück zu den eben erwähnten Plausibilitätsproblemen.

2 Das Problem: Wer oder was ist „zentral“?

Das Betweenness-Modell ist ein konzeptionell einleuchtendes und auch theoretisch fundiertes Konzept, das in den meisten Fällen erstaunlich gute Resultate liefert. Doch hin und wieder wundert man sich auch: der zentralste Akteur in den deutschen Sozialwissenschaften – ein Diplom-Ingenieur am Fraunhofer-Institut für Produktionstechnik und Automatisierung?⁵ Auch in anderen Disziplinen kennt man diese (gelegentliche) Verwunderung: Anchorage – der zentralste Flughafen im weltweiten Flugverbindungsnetzwerk (Guimerà et al. 2005)?

Handelt es sich hierbei (nur) um statistische Ausreißer oder (doch) um eine konzeptionelle Schwäche des Zentralitätsmodells? Diese Frage kann letztlich natürlich nur im jeweiligen Anwendungskontext entschieden werden. Außer Frage steht allerdings, dass Betweenness ein sehr mächtiges Konzept ist, da es Zentralität auf der Basis (indirekter) triangularer Beziehungen über die gesamte Netzwerkstruktur hinweg berechnet. Hierin liegt sowohl die Mächtigkeit, zugleich aber auch die interpretative Schwierigkeit dieses Konzepts. Denn die Grundannahme des Betweenness-Modells ist, dass ein (zentraler) Akteur

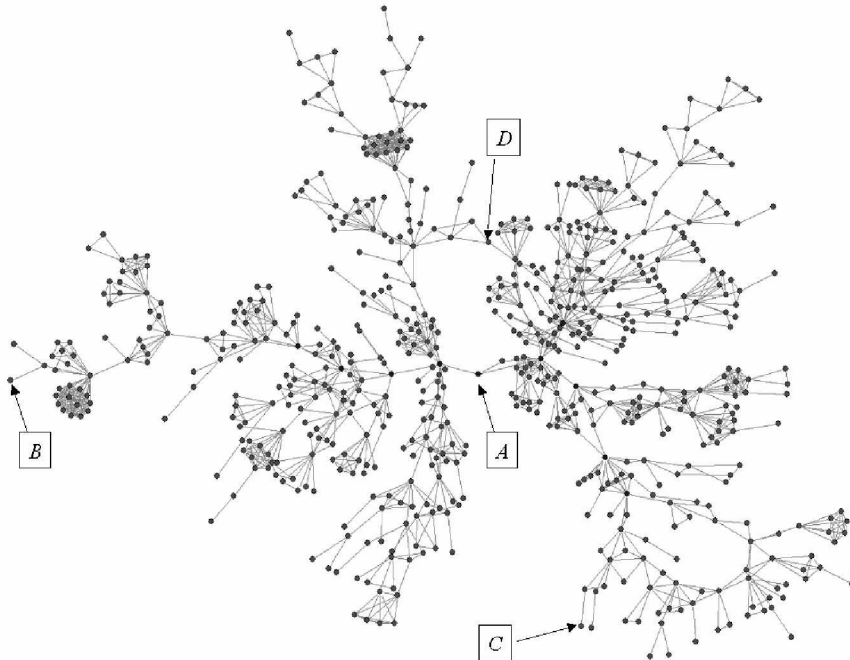
³ www.infoconnex.de; vgl. auch das Daffodil-System (www.daffodil.de).

⁴ Precision ist ein Maß aus dem Information Retrieval, dass über die Rate Auskunft gibt, mit der das Ranking relevante Treffer liefert.

⁵ Zentralitätsanalyse auf Basis der Datenbank SOLIS der GESIS (<http://www.gesis.org/Information/SOLIS/>), August 2007. Das bei einer solchen Analyse auch Datenbankeffekte eine Rolle spielen, sei hier nur angemerkt.

die Interaktion auch zwischen, graphentheoretisch betrachtet, weit von einander entfernten Akteuren kontrolliert. Die Frage ist nur, ob diese Grundannahme wirklich der sozialen Realität in einem realen Kommunikations- und Kooperationsnetzwerk entspricht. In dem Autorennetzwerk in Abbildung 1 z.B. hat der Knoten *A* die dritt-höchste Betweenness.⁶ Doch kontrolliert *A* wirklich den Informationstransfer z.B. zwischen *B* und *C*? Und wenn ja, wie wäre diese Beziehung dann zu interpretieren, insbesondere im Unterschied zu anderen (viel kürzeren) Beziehungen, die *A* kontrolliert?

Abbildung 1: Autorennetzwerk zum Thema „Kindheit“⁷



Um es vorwegzunehmen: Es geht in diesem Beitrag weder um die Wahl der „richtigen“ Transaktionspfade in Graphen noch um ein Betweenness-Maß, dass mit Gewichtungen oder Filtern (z.B. Schwellwerten für Pfadlängen) arbeitet, und schon gar nicht um eine neue Definition von „Zentralität“ und ein neues (weiteres) Zentralitätsmaß. Es geht vielmehr um den Verdacht, dass die oben skizzierten Plausibilitäts- und Interpretationsprobleme eher etwas mit unserem intuitiven Verständnis von „Netzwerk“ und „Zentralität“ zu tun haben. Kann man vor dem Hintergrund unserer realen Erfahrung von sozialen „Netzwerken“ überhaupt sagen, *B* und *C* seien „vernetzt“? Rein graphentheoretisch betrachtet sind sie es natürlich – dies will der Beitrag grundsätzlich nicht bestreiten. Die Frage (und Schwierigkeit) ist nur, ob diese Beziehung im Kontext der sozialen Realität, in die sie eingebettet ist, und

⁶ Ein für dieses Forschungsfeld inhaltlich unplausibles Ergebnis; gleiches gilt für den Knoten *D*.

⁷ Autorennetzwerk zum Forschungsfeld „Kindheit“ (ab 1997) auf Basis der Datenbank SOLIS. Evaluationsstudie in Kooperation mit dem Deutschen Jugendinstitut, April 2007 (unveröffentlicht). Die Grafik wurde mit dem Programm Pajek (<http://vlado.fmf.uni-lj.si/pub/networks/pajek/>) generiert.

insbesondere im Hinblick auf die Zentralitätsaussage, für die sie herangezogen wird (z.B. die positionale Evaluation des Knotens *A*), überhaupt noch eine inhaltliche Bedeutung hat. Dieses interpretative Defizit, das auch von der traditionellen Netzwerkanalyse eingeräumt wird (Borgatti/ Everett 2006), verleitet zu der ketzerischen Frage, ob es sich bei Abbildung 1 überhaupt um ein *reales* Netzwerk handelt oder „nur“ um eine netzwerkartige Visualisierung der Beziehungsdaten. Jede netzwerkanalytische Anwendung sieht sich der latenten Gefahr gegenüber, Artefakte statt veritable Interaktionsmodelle der betrachteten Gesellschaft zu generieren und den Netzwerkbegriff somit nur noch metaphorisch zu verwenden. Umberto Eco beschreibt dieses Problem in seinem großen Verschwörungsroman „Das Foucaultsche Pendel“ mit folgenden treffenden Worten:

„... denn wenn man Zusammenhänge finden will, findet man immer welche, Zusammenhänge zwischen allem und jedem, die Welt explodiert zu einem wirbelnden Netz von Verwandtschaften, in dem alles auf alles verweist und alles alles erklärt“. (dtv-Ausgabe, 1992, S. 600)

Friedkin (1983) konnte zeigen, dass Netzwerkmitglieder, die mehr als zwei Links von einander entfernt sind, einander nicht mehr wahrnehmen (vgl. auch Lazer 2001). Dieses von Friedkin als *horizon of observability* bezeichnete Phänomen wird auch von einigen Small-World-Experimenten bestätigt (vgl. Adamic/ Adar 2005). Vor diesem Hintergrund muss man die Frage nach der Interpretationsfähigkeit großer Netzwerke eindeutig verneinen. Die Validität von Zentralitätsaussagen hat also auch etwas mit dem Netzwerkbegriff zu tun, der unserer Anwendung zugrunde liegt. Welche Schlussfolgerungen ziehen wir daraus für unser Zentralitätsproblem? Schauen wir uns die drei Standardmaße Degree, Closeness und Betweenness vor dem Hintergrund dieses interpretativen Defizits noch einmal genauer an.

3 Empirie der Zentralitätsanalyse: Anspruch und Anomalien

Zentralitätsmaße sind in der (sozialen) Netzwerkanalyse standardmäßig mit ganz bestimmten konzeptionellen Grundannahmen verknüpft (vgl. Wasserman/ Faust 1994): Degree misst die soziale Aktivität eines Akteurs anhand der Anzahl seiner direkten Verbindungen zu anderen Akteuren im Netzwerk. Closeness erfasst die strukturelle Unabhängigkeit bzw. Effizienz eines Akteurs anhand seiner Pfaddistanz (Nähe) zu allen anderen Knoten des Netzwerkes. Betweenness misst den Grad der strukturellen Abhängigkeit vom betrachteten Akteur bzw. das Kontrollpotential eines Akteurs anhand der Anzahl der kürzesten Verbindungen zwischen Knoten im Netzwerk, die über den betrachteten Akteur laufen. Alle drei Standardmaße repräsentieren eine bestimmte und durchaus plausible Vorstellung von Zentralität, die auf folgende konzeptionelle Grundkategorien zurückgeführt werden kann: *Aktivität* (Degree), *Effizienz* (Closeness), *Kontrolle* (Betweenness). Betweenness ist dabei das in der Netzwerk-Community am meisten beachtete Zentralitätsmodell. Doch warum liefert gerade dieses Maß gelegentlich qualitativ höchst unplausible Ergebnisse?

Sieht man sich die Fälle von inhaltlich unplausibel hoher Zentralität genauer an, wie z.B. die hohe Betweenness des Knotens *A* in Abbildung 1, dann fällt empirisch in der Regel eine mangelnde Korrespondenz von mindestens zwei der o.g. Kategorien auf. Im Falle des

Knotens *A* von Abbildung 1 ist es die fehlende Korrespondenz von *Kontrolle* und *Aktivität*⁸. Eine erste explorative Untersuchung unplausibler Rankings in infoconnex ergab tatsächlich, dass Plausibilitätsprobleme in der Regel mit Statusinkonsistenzen einhergehen, d.h. wo die Resultate verschiedener Zentralitätsmaße bei einem Akteur in einem deutlichen Missverhältnis zueinander stehen (z.B., wie im Falle des Knotens *A*, ein schwacher Degree bei gleichzeitig hoher Betweenness). In einem solchen Fall liegt eine *Zentralitätsanomalie*⁹ vor, die den Wert der Zentralitätsaussage mindert oder sogar ganz in Frage stellt. Aus dieser Sichtweise lässt sich folgende paarweise Anomalie-Systematik für die Standardmaße Degree, Closeness und Betweenness ableiten:

Betweenness-Anomalien:

- Degree-Betweenness-Anomalie¹⁰ (hohe Betweenness, geringer Degree): Missverhältnis zwischen Kontrolle und Aktivität, verursacht durch Abstraktion des Betweenness-Formalismus von der Stärke der lokalen Vernetzung; typische Fälle sind „cutpoints“, die über nur wenige Ko-Akteure zwei Cluster miteinander verbinden.
- Closeness-Betweenness-Anomalie (hohe Betweenness, geringe Closeness): Missverhältnis zwischen Kontrolle und Effizienz, verursacht durch Abstraktion des Betweenness-Formalismus“ von Pfaddistanzen; kritische Akteure sind „Broker“, die an der Peripherie eines Netzwerkes angesiedelt sind.

Degree-Anomalien:

- Betweenness-Degree-Anomalie (hoher Degree, geringe Betweenness): Missverhältnis zwischen Aktivität und Kontrolle, verursacht durch Abstraktion des Degree-Formalismus' vom Grad der strukturellen Abhängigkeit der Ko-Akteure vom betrachteten Akteur; typische Kandidaten sind stark vernetzte Cliquenmitglieder.
- Closeness-Degree-Anomalie (hoher Degree, geringe Closeness): Missverhältnis zwischen Aktivität und Effizienz, verursacht durch Abstraktion des Degree-Formalismus' vom Grad der geodätischen Nähe des betrachteten Akteurs zu anderen Akteuren im Netzwerk; kritische Fälle sind stark vernetzte Akteure an der Peripherie eines Netzwerkes.

Closeness-Anomalien:

- Betweenness-Closeness-Anomalie (hohe Closeness, geringe Betweenness): Missverhältnis zwischen Effizienz und Kontrolle, verursacht durch Abstraktion des Closeness-Formalismus' vom Grad der strukturellen Abhängigkeit vom betrachteten Akteur; kritische Akteure sind Mitglieder einer zentralen Clique.
- Degree-Closeness-Anomalie (hohe Closeness, geringer Degree): Missverhältnis zwischen Effizienz und Aktivität, verursacht durch Abstraktion des Closeness-Formalismus' vom Grad der lokalen Vernetzung; typischer Kandidat ist der „hangers-on“ mit nur einer Verbindung zu einem zentralen Akteur „in der Mitte“ des Netzwerkes.

Von diesen Anomalien ist die Degree-Betweenness-Anomalie (DBA) diejenige, die auch für die eingangs skizzierten Plausibilitätsprobleme verantwortlich ist. Auch der Knoten *A* in

⁸ Gleiches gilt übrigens auch für den Fall des zentralsten deutschen Sozialwissenschaftlers in SOLIS und für den Fall des zentralsten Flughafens.

⁹ Der Anomaliebegriff wird erstmals von Guimerà et al. (2005) verwendet, um Inkonsistenzen zwischen Degree- und Betweenness-Werten zu beschreiben.

¹⁰ Guimerà et al. 2005.

Abbildung 1 ist ein typischer DBA-Knoten. Dieser Befund korrespondiert mit Ergebnissen empirischer Untersuchungen zu Zentralitätsmaßen, welche eine starke Zufallsanfälligkeit der Betweenness-Zentralität festgestellt haben (Trappmann et al. 2005, Kim/ Jeong 2007). Bei Auftreten einer Betweenness-Anomalie dürfte dieses Sensibilitätsproblem umso schwerwiegender ins Gewicht fallen, weshalb man mit der Interpretation von Betweenness-Zentralitätswerten, obwohl diese grundsätzlich wertvolle Erkenntnisse über die Position von Akteuren im Netzwerk vermitteln, gerade bei anomalen Werten besonderes vorsichtig sein sollte. Das Interessante ist nun, dass Zentralitätsanomalien, und hier insbesondere die DBA, nicht unbedingt mit den Erkenntnissen der Netzwerkforschung über die Struktureigenschaften sozialer Netzwerke in Einklang zu bringen sind. Betrachten wir uns deshalb diese Erkenntnisse einmal vor dem Hintergrund der oben skizzierten Anomalien.

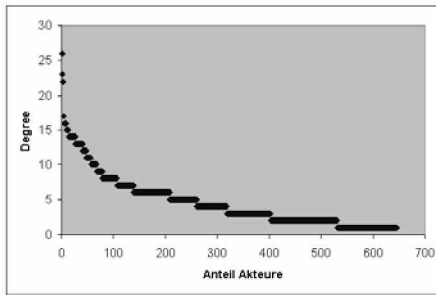
4 Netzwerkstruktur und Zentralität: Die (gerne beachteten und doch) vergessenen Communities

Eine fundamentales Strukturmerkmal komplexer vernetzter Real-World-Systeme ist die sog. „Small-World“-Eigenschaft (vgl. Watts 1999; Newman 2006). In einem Small-World-Netzwerk sind die Knoten – im Unterschied zu Zufallsgraphen – über relativ kurze Wege miteinander verbunden. Darüber hinaus haben die Knoten einen hohen Grad an Cliquenbildung, d.h. die Tendenz, lokale Communities auszubilden, die intern stärker vernetzt sind als mit dem Rest des Netzwerkes. Die strukturelle Natur sozialer Netzwerke zeichnet sich also durch das Vorhandensein relativ kurzer Verbindungen von einem Knoten zum anderen als auch durch die Existenz in sich dicht gestrickter Subsysteme aus, die eine gewisse Homogenität¹¹ unter ihren Mitgliedern repräsentieren und nur locker mit anderen Communities verbunden sind. Ein weiterer elementarer Faktor, der Real-World-Netzwerke von Zufallsgraphen unterscheidet, ist die Degree-Verteilung, d.h. die Verteilung der Verbindungen auf die Knoten, deren Verlauf in Real-World-Netzwerken dadurch charakterisiert ist, dass einige wenige Knoten eine deutlich höhere Anzahl Verbindungen auf sich vereinigen als die meisten anderen Knoten. Diese (auch als „skalen-frei“ bezeichneten) Netzwerke haben also eine starke Tendenz zur Zentralisierung (Albert/ Barabási 2002).

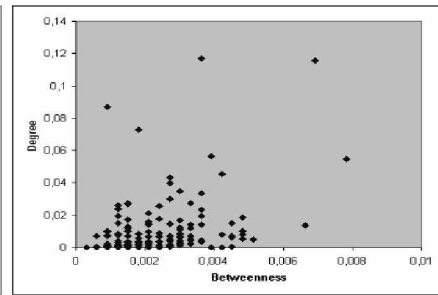
Auch unser Beispielnetzwerk von Abbildung 1 ist ein Small-World-Netzwerk mit einer starken Tendenz zur Zentralisierung: die charakteristische Pfadlänge ist mit 9,4 aufgrund der starken Fragmentierung des Netzwerkes zwar relativ hoch, entspricht aber immer noch dem typischen Verhalten von Small-World-Netzwerken; der Clustering-Koeffizient, der über die lokale Vernetzungsrate Auskunft gibt, ist mit 61 % deutlich ausgeprägt; auch die Degree-Verteilung zeigt das typische Verhalten skalen-freier Netzwerke (s. Abb. 2a). Schaut man sich aber nun für dieses Netzwerk an, inwieweit Degree und Betweenness miteinander korrelieren, dann zeigt sich – im Unterschied zu Zufallsgraphen, wo Degree und Betweenness stark miteinander korrelieren (vgl. Guimerà et al. 2005) – mit einem Korrelationskoeffizienten von 0,38 ein relativ gemäßigter statistischer Zusammenhang (s. Abb. 2b). Die Grund dafür ist, dass es einige Akteure mit einem kleinen Degree, aber einer sehr hohen Betweenness gibt, und umgekehrt viele Akteure mit einem hohen Degree, aber relativ kleinen Betweenness. Nur ein einziger Akteur hat zugleich eine hohe Betweenness und einen hohen Degree. Das Netzwerk zeichnet sich also durch einige ausgeprägte DBAs aus.

¹¹ Siehe das Homophilie-Paradigma, vgl. Lazer 2001.

Abbildung 2: a) Degree-Verteilung



b) Degree-Betweenness-Korrelation



Betrachtet man sich die DBA nun vor dem Hintergrund der oben skizzierten Struktureigenschaften vernetzter Systeme, dann fällt auf, dass sie in einem ambivalenten Verhältnis zur Small-World-Struktur sozialer Netzwerke steht: Sie ist einerseits mitverantwortlich für kurze Entfernungen im Netzwerk („Small-World“), repräsentieren doch gerade DBA-Knoten i.d.R. periphere, Communities verbindende Akteure („weak ties“). Andererseits läuft die DBA bzw. das sie produzierende Zentralitätsmodell (Betweenness) der Community-Struktur sozialer Netzwerke zuwider. Denn was die DBA ausmacht, ist ja gerade der Widerspruch zwischen peripherem Status eines DBA-Knotens innerhalb seiner Community („strong ties“) und seiner hohen Betweenness im ganzen Netzwerk (vgl. auch Borgatti/ Everett 2006). Da aber Community-Struktur auch unserer alltäglichen Wahrnehmung vernetzter sozialer Systeme entspricht (siehe „network horizon“), haben wir nun auch eine strukturelle Erklärung für den mit einer DBA einhergehenden Plausibilitätsverlust der Zentralitätsaussage gefunden, nämlich die mit einer DBA verbundene Gegenläufigkeit zur Community-Eigenschaft sozialer Netzwerke. An dieser Stelle wird zugleich eine gewisse Widersprüchlichkeit zwischen der Wahrnehmung von Community-Grenzen als Netzwerk-grenzen („horizon of observability“) und dem Netzwerkbegriff der traditionellen Netzwerkanalyse deutlich, der von einer globalen Gleichrangigkeit aller (noch so indirekten) Beziehungen in einem Netzwerk ausgeht. Vor dem Hintergrund der strukturellen Beschaffenheit sozialer Netzwerke wäre ein „Netzwerk“ allerdings weniger als ein Verbund miteinander interagierender Individuen (die bisherige Sichtweise) zu verstehen, sondern vielmehr als ein heterogener Verbund in sich relativ homogener Communities.

Überraschend ist nun, wie Guimerà et al. (2005) gezeigt haben, dass die DBA durch die Existenz verschiedener distinkter Communities erklärt werden kann. Wenn wir noch einmal einen Blick auf das Autorennetzwerk von Abbildung 1 werfen, sehen wir zwei größere Regionen mit teilweise hoher interner Dichte (Communities), die nur über zwei einzelne Knoten miteinander verbunden sind, nämlich durch den Knoten *D* im oberen Bereich der Grafik und durch unseren Knoten *A* in der Mitte der Abbildung. Eine Verbindung zur jeweils anderen Community existiert nur über die über *A* und *D* laufenden Verbindungen. Dieses Faktum erklärt die hohe Betweenness insbesondere von *A*, weil dieser die Hauptverbindung zwischen den beiden Communities ist. Die Existenz eines dezidierten DBA-Knotens (hier der Knoten *A*), d.h. eines Knotens mit anormaler Zentralität, steht also in enger Relation zu der Existenz von Communities. Community-Struktur ist andererseits aber zugleich genau das Merkmal sozialer Netzwerke, dass in einem globalen Zentralitätsmodell

wie Betweenness nicht vorkommt – mit dem aus qualitativer Sicht problematischen Seiteneffekt, dass eine (u.U. zufällige) Verbindung zwischen peripheren Akteuren verschiedener Communities zu einer zentralen Verbindung für das ganze Netzwerk „hochgerechnet“ wird.

Die Kernthese dieses Papers ist somit, dass in den globalen Zentralitätsmodellen der traditionellen Netzwerkanalyse (wie Betweenness) eines der fundamentalsten Strukturmerkmale sozialer Netzwerke nicht berücksichtigt ist: nämlich die Tendenz der Netzwerke, Communities zu bilden. Zentralitätsanomalien sind somit strukturinduzierte Statusinkonsistenzen, die diese Inkompatibilität zwischen der strukturellen Natur des Netzwerkes und der konzeptionellen Natur des Zentralitätsmodells reflektieren.

Dies mag nicht für jede Anwendung ein Problem sein. Bei der Erhebung und Analyse eines Terroristennetzwerkes, wo der Nachweis der Existenz einer Verbindung an sich das zentrale Forschungsziel ist, wird auch das traditionelle globale Netzwerk- und Zentralitätsmodell ein legitimer und ausreichender Analyseansatz sein. Dort wird die Entlarvung eines DBA-Knotens vermutlich sogar die entscheidende Entdeckung der Anwendung sein. In jedem Anwendungskontext aber, wo die Existenz von Communities ein primäres Strukturmerkmal des Netzwerkes ist (wie es bei Autorennetzwerken sicherlich der Fall ist) und wo darüber hinaus die Relevanz einer Verbindung womöglich auch noch mit einer gewissen Vagheit behaftet ist¹², sollte ein Netzwerkverständnis zugrunde gelegt werden, der die Community-Struktur sozialer Netzwerke angemessen berücksichtigt. Im Folgenden wird gezeigt, wie ein Zentralitätsmodell auf der Basis eines „engeren“, an der Community-Struktur sozialer Netzwerke orientierten Netzwerkbegriffs konzeptionell aussehen könnte.

5 Alternative: Ein rekursives community-orientiertes Metamodell von Zentralität auf der Basis eines „engeren“ Netzwerkbegriffs

Für den Umgang mit Zentralitätsanomalien bieten sich mehrere Strategien an: (1) Eine in der traditionellen Netzwerkanalyse gebräuchliche Variante ist, Statusinkonsistenzen nicht als Anomalien wahrzunehmen, sondern als unterschiedliche *Rollen*, die Akteure in der Netzwerkstruktur einnehmen (vgl. Borgatti/ Everett 2006). Das Problem dieser Sichtweise ist jedoch, dass sie Zentralitätswerte, die aus qualitativer Sicht unplausibel sind, nicht erklären kann. Allerdings räumt auch die traditionelle Netzwerkanalyse ein, dass „mediale“ Zentralitätswerte wie Betweenness ohne Kenntnis der kohäsiven Struktur des Netzwerkes schwierig zu interpretieren und „radiale“ Zentralitätskonzepte wie Closeness eigentlich nur bei stark zentralisierten Netzwerken sinnvoll seien (Borgatti/ Everett 2006). (2) Eine mögliche Strategie zur Vermeidung von Zentralitätsanomalien ist die geschickte Verwendung von Schwellwerten oder Gewichtungen bei Pfadlängen (*k-betweenness*, *length-scaled-betweenness*, siehe Borgatti/ Everett 2006). Neben dem Problem des arbiträren Charakters von Schwellwerten hat diese Strategie aber ebenfalls die Schwäche, grundsätzlich Strukturmerkmale sozialer Netzwerke (wie Community-Struktur) zu ignorieren. (3) Eine weitaus erfolgversprechendere Vermeidungsstrategie besteht darin, Anomalieknoten auszufiltern. Diese Variante hat den Vorzug, dass Knoten mit anomalen Zentralitätswerten von der weiteren Betrachtung, z.B. einem nachträglichen Ranking, ausgeklammert werden könnten, um

¹² Dies dürfte v.a. in Anwendungen der Fall sein, wo Netzwerke über Zeit entstehen und verzerrende Effekte durch singuläre Beziehungsereignisse entstehen können (vgl. auch den Beitrag von Trier und Bobrik zu Sudden-Burst-Effekten in diesem Band).

Plausibilitätsverluste möglichst zu vermeiden. Ein mathematisches Modell für die Identifikation von Anomalieknoten steht allerdings noch aus. Außerdem gilt auch für diese Variante, dass sie auf dem konventionellen, Community-Struktur ignorierenden Zentralitätsmodell aufsetzt.

Eine echte Alternative (4) wäre ein Zentralitätsmodell, welches von einem „engeren“ Netzwerkbegriff ausgeht. „Eng“ bedeutet hier die Absage an das (globale) „Vernetzungs“-Paradigma der traditionellen Netzwerkanalyse, das davon ausgeht, dass alle Individuen auf die gleiche Weise in ein globales Beziehungsgeflecht „eingebettet“ sind und daher für alle Beziehungen in einem Netzwerk, seien sie auch noch so indirekt, eine gleichförmige Betrachtungsweise propagiert. Das Paper plädiert vielmehr für ein Verständnis von „Vernetzung“, das sowohl Community-Grenzen als auch die *Observability* von Beziehungen im Sinne des Friedkin'schen *network horizon* angemessen berücksichtigt. Die beiden Grundprämissen eines in diesem Sinne „engeren“ Netzwerkbegriffs sind somit, aus struktureller Perspektive, die Existenz von Community-Struktur und, aus der kognitiven Wahrnehmungsperspektive des Individuums, die Berücksichtigung der Tatsache, dass „Netzwerke“ nur innerhalb eines bestimmten (kleinen) Radius wahrgenommen werden („horizon of observability“).

Diese Sichtweise hat den sowohl theoretisch als auch für die praktische Anwendung von Netzwerkanalysen bedeutenden Vorzug, dass beide Perspektiven in einer symbiotischen Wechselbeziehung zueinander stehen: Community-Grenzen beschreiben die Grenzen des „eigenen“, wahrgenommen Netzwerkes; die Wahrnehmung der eigenen Community impliziert zugleich aber auch, nämlich auf der Meta-Ebene der Communities, die Wahrnehmung von Community-Nachbarschaften, die sich über Community-externe Beziehungen konstituieren. Ein „engeres“ Verständnis von Netzwerk bedeutet daher keine Einschränkung von Netzwerkanalyse allein auf Communities – im Gegenteil: die Berücksichtigung sowohl von Community-Struktur als auch von *Observability* lässt auch die Analyse von Beziehungen zu, die über Community-Grenzen hinausgehen, da der *network horizon* ohne weiteres von der Ebene des Individuums auf die Ebene der Community angehoben werden kann, nämlich als Wahrnehmung von Community-Nachbarschaften. Gleichzeitig bietet die Community-Struktur eines Netzwerkes selbst Analyseperspektiven, nämlich die Analyse von Zentralität auf der Basis des Community-Netzwerkes.

Das Paper schlägt daher ein community-orientiertes Metamodell von Zentralität vor, dass sowohl lokale Clusterung als auch *Observability* in sich vereint, indem es Zentralität auf der Basis der Community-Struktur des Netzwerkes evaluiert und somit eine semantisch zuverlässigere Interpretierbarkeit von Zentralitätsaussagen verspricht. *Observability* fungiert dabei als Parameter, der die maximale Tiefe einer Community definiert¹³. Kernstück dieses Ansatzes ist die explizite Unterscheidung zwischen den Entity-Ebenen Akteur (Knoten) und Community innerhalb eines Zentralitätsmodells, so dass auch eine rekursive Anwendung des Zentralitätsmodells möglich ist. Diesem Ansatz liegt ein hierarchisches Netzwerkmodell zugrunde, das davon ausgeht, dass Individuen zu Gruppen (Communities) gehören, die ihrerseits wiederum in größere Gruppen (Meta-Communities) eingebettet sind (vgl. Adamic/ Adar 2005). Auf der Ebene der Meta-Communities wären die Communities

¹³ Ein dem Friedkin'schen *network horizon* gemäßer Standardwert wäre 2, so dass alle Mitglieder einer Community nicht mehr als zwei Schritte voneinander entfernt wären. Es ist allerdings zu vermuten, dass dieser Wert zu tendenziell sehr kleinteiligen „Communities“ führen würde, so dass in den meisten Fällen ein größerer *Observability*-Wert ratsam erscheint.

demnach Cluster in sich stärker vernetzter kleinerer Communities, die dann die Akteurs-ebene repräsentieren.

Der entscheidende Unterschied zum traditionellen Zentralitätsansatz besteht darin, dass Zentralität nicht auf Basis der ganzen horizontalen Ausdehnung des Netzwerkes berechnet wird, sondern an Community-Grenzen halt macht.¹⁴ Auf diese Weise werden Zentralitätsanomalien zwar nicht prinzipiell vermieden, aber auf die Communities eingeschränkt und verhindert, dass sie auf das ganze Netzwerk übertragen werden. Anomalien würden, sofern sie auftreten, gewissermaßen in „ihrer“ Community „gekapselt“. Zusätzlich wird, im Sinne einer zweidimensionalen Vorstellung von Zentralität, auch die vertikale Struktur des Netzwerkes evaluiert, indem auch eine positionale Analyse der Communities durchgeführt wird. Ein solches die Community-Struktur sozialer Netzwerke berücksichtigendes Zentralitätskonzept bietet eine ganze Reihe interessanter (z.T. neuer) Anwendungsperspektiven:

- *Internal Centrality* (Akteurs-Ebene): Zentralität eines Knotens innerhalb seiner Community (internal degree¹⁵, internal closeness, internal betweenness);
- *Community Centrality*¹⁶ (Community-Ebene): Zentralität innerhalb des Community-Netzwerkes, d.h. die Communities werden als Knoten eines Netzwerkes aufgefasst und Zentralität wird auf Basis des group reduced graph (Everett/ Borgatti 1999) mit den gängigen Methoden berechnet (community degree, intra-community degree¹⁷, community closeness, community betweenness); eine in diesem Zusammenhang interessante Variante ist group centrality (Everett/ Borgatti 1999), d.h. die Zentralität einer gegebenen Community innerhalb des initialen (Akteurs-)Netzwerkes;
- *External Centrality* (Akteursebene): Zahl der community-externen Verbindungen eines Akteurs; hierfür sind zwei Varianten zu unterscheiden: die Zahl der Verbindungen mit community-externen Akteuren (external vertex degree) und die Zahl der Verbindungen zu anderen Communities (external community degree)¹⁸;
- *Nested Vertex Centrality* (Akteursebene): die Zentralität eines Akteurs innerhalb seiner Community (internal centrality) wird mit der Zentralität der jeweiligen Community innerhalb des Community-Netzwerkes (community centrality) verknüpft; die Zentralität eines Akteurs erhöht sich also um die Zentralität der Community, so dass Akteure zentraler Communities zentraler sind als Akteure, die eine vergleichbare internal centrality haben, aber in peripheren Communities lokalisiert sind;¹⁹
- *Nested Community Centrality* (Community-Ebene): community centrality wird erhöht um Eigenschaften des community-internen Netzwerkes (Zentralisierung, Dichte);

¹⁴ Dieses Vorgehen bietet auch einen vielversprechenden Lösungsansatz für das Komponentenproblem, d.h. der Benachteiligung von Mitgliedern kleinerer Komponenten in globalen Zentralitätsmodellen.

¹⁵ Entspricht dem within-community-degree bei Guimerà et al. (2005); vgl. auch Newman (2006).

¹⁶ Bei Newman (2006) ist „Community Centrality“ ein akteursbezogener Index, der den Beitrag eines Akteurs zur Modularität des Netzwerkes misst. Dies scheint jedoch mehr auf eine Dezentralitäts- als auf eine Zentralitätsaussage hinauszulaufen, weshalb hier vorgeschlagen wird, „Community Centrality“ als Maß für die Zentralität einer Community im Community-Netzwerk zu betrachten.

¹⁷ Zahl der Community-Mitglieder (im Unterschied zum community degree, der ein inter-community degree ist).

¹⁸ Guimerà et al. (2005) schlagen einen „Participation Coefficient“ vor, der die Verteilung externer Verbindungen auf die Communities misst.

¹⁹ Für eine Formalisierung dieses Konzepts ist noch weitere Forschungsarbeit zu leisten.

- *Hubs Centrality* (Akteursebene): Zentralität der Community-Hubs in einem auf die lokalen Hubs und deren Verbindungen zu Nachbar-Communities reduzierten Netzwerk,²⁰ ein solches Hubs-Netzwerk könnte auch eine vielversprechende Basis für High-Degree-Suchstrategien sein.²¹

Zu beachten ist, dass die vorgeschlagenen Maße nicht unterschiedliche Anwendungen von Zentralitätsanalysen, sondern unterschiedliche Analyseebenen innerhalb *einer* Zentralitätsanalyse repräsentieren. Dies bedeutet, dass sie rekursiv auf die jeweils höhere Community-Ebene angewendet werden können, d.h. auf der Ebene der Meta-Communities, Meta-Meta-Communities usw. Das vorgeschlagene Zentralitätsmodell iteriert also rekursiv durch die Struktur der verschiedenen Community-Ebenen, wobei auf jeder Ebene die konkreten ebenen-bezogenen Zentralitätskonzepte (internal centrality, community centrality etc.) Anwendung finden.

Diese Anwendungsszenarien eines rekursiven, community-orientierten Zentralitätsmodells versprechen ein erhebliches Innovationspotential für die Zentralitätsanalyse und neue Einsichten in die positionale Evaluation sowohl von Individuen als auch von Communities. Die Voraussetzung hierfür ist allerdings jeweils eine Community-Analyse, die auf jeder Ebene der Zentralitätsanalyse zunächst durchgeführt werden muss.²² Dieses Vorgehen impliziert allerdings eine fundamentale Veränderung der Perspektive auf das Ausgangsmaterial: die traditionelle Adjacency-Matrix ist in diesem Modell nicht mehr die Basis der Zentralitätsanalyse, sondern (nur noch) Input für die Community-Analyse. Durch die Community-Analyse werden die Ausgangsdaten zunächst in eine Normalform überführt, welche die (rekursiv abgestufte) Community-Struktur des Netzwerkes abbildet. Diese *Community-Normalform* (CSNF²³) bildet die Grundlage für alle nachfolgenden Zentralitätsanalysen. Dem Phasenmodell im Knowledge-Discovery-Prozess vergleichbar wird die Community-Analyse somit zu einer Datenbereinigungskomponente, die aus den rohen Beziehungsdaten, die in der traditionellen Netzwerkanalyse *das* Netzwerk repräsentieren, erst ein Netzwerk in dem hier vorgeschlagenen „engeren“ Sinne macht. Die Entwicklung eines geeigneten Formalismus für ein solches Verfahren steht allerdings noch aus. Eine CSNF der Netzwerkdaten würde aber die soziale Realität vieler Interaktionsstrukturen, auch unter dem Gesichtspunkt des *network horizon*, sicherlich angemessener abbilden als dies die traditionelle Adjacency-Matrix zu leisten imstande ist.

6 Literatur

- Adamic, L.A.; Adar, E. (2005): How to search a social network, *Social Networks* 27, 3: 187-203.
 Albert, R.; Barabási, A.-L. (2002): Statistical mechanics of complex networks, *Reviews of Modern Physics* 74: 47-97.

²⁰ Eine interessante Fragestellung für weiterführende Analysen wäre, ob es typischerweise die Hubs sind, die Verbindungen zu anderen Communities unterhalten oder eher die peripheren Akteure einer Community.

²¹ Eine solche Strategie scheint gerade in Netzwerken mit exponentiellen Degree-Verteilungen effektiv zu sein (s. Adamic/Adar (2005); vgl. auch Mutschke (2004)).

²² Für Community-Analysen wurden in den letzten Jahren eine Reihe performanter Algorithmen vorgestellt. Stellvertretend für viele seien hier die Arbeiten von Newman (2006) genannt.

²³ Community Structure Normal Form

- Beaver, D. de B. (2004): Does collaborative research have greater epistemic authority? *Scientometrics* 60, 3: 309-408.
- Borgatti, S. P.; Everett, M. G. (2006): A Graph-theoretic perspective on centrality, *Social Networks* 28, 4: 466-484.
- Brin, S.; Page, L. (1998): The Anatomy of a Large-Scale Hypertextual Web Search Engine, *Computer Networks and ISDN Systems* 30: 107-117.
- Everett, M. G.; Borgatti, S. P. (1999): The centrality of groups and classes, *Journal of Mathematical Sociology* 23, 3: 181-201.
- Guimerà, R.; Mossa, S.; Turtschi, A.; Amaral, L. A. N. (2005): The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles, *PNAS* 102, 22: 7794-7799.
- Friedkin, N. E. (1983): Horizons of Observability and Limits of Informal Control in Organizations, *Social Forces* 62, 1: 54-77.
- Lazer, D. (2001): The Co-Evolution of Individual and Network, *Journal of Mathematical Sociology* 25, 1: 69-108.
- Kim, P.-J.; Jeong H. (2007): Reliability of rank order in sampled networks, *Eur. Phys. J. B* 55: 109-114.
- Mutschke, P. (2004): Autorennetzwerke: Netzwerkanalyse als Mehrwertdienst für Informationssysteme, S. 141-162, in: Bekavac, B. et al. (Hrsg.): *Information zwischen Kultur und Marktwirtschaft. Proceedings des 9. Internationalen Symposiums für Informationswissenschaft (ISI 2004)*, Chur. Konstanz: UVK (Schriften zur Informationswissenschaft, Band 42).
- Mutschke, P.; Stahl, M. (2005): Kooperationsnetzwerke und Akteurszentralität im Forschungsfeld Bildung, S. 121-140, in: Engel, U. (Hrsg.): *Bildung und soziale Ungleichheit: methodologische und strukturelle Analysen*. Bonn: IZ Sozialwissenschaften. (Tagungsberichte; Bd. 9).
- Newman, M. E. J. (2006): Finding community structure in networks using the eigenvectors of matrices. *Phys. Rev. E* 74, 036104.
- Trappmann, M.; Hummell H. J.; Sodeur, W. (2005): *Strukturanalyse sozialer Netzwerke. Konzepte, Modelle, Methoden*. Wiesbaden: VS.
- Watts, D. J. (1999): *Small Worlds*. Princeton: Princeton University Press.
- Wasserman, S.; Faust, K. (1994): *Social Network Analysis. Methods and Applications*: Cambridge University Press (Structural Analysis in the Social Sciences, 8).